DOCUMENT RESUME

ED 129 835 TM 005 233

AUTHOR Bergsten, Jane W.
TITLE Sample Design and Data Collection: National Study of the Upward Bound Program.
INSTITUTION Research Triangle Inst., Durham, N.C.
PUB DATE Apr 76
NOTE 31p.; Paper presented at the Annual Meeting of the American Educational Research Association (60th, San Francisco, California, April 19-23, 1976); For related documents, see ED 121 994 and ED 121 995

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Comparative Analysis; *Compensatory Education Programs; *Data Collection; Dropout Rate; High School Students; Low Income Groups; Program Effectiveness; *Program Evaluation; Questionnaires; *Research Design; *Sampling; School Holding Power; *Secondary Education
IDENTIFIERS *Upward Bound

ABSTRACT

The Research Triangle conducted an evaluation study of the Upward Bound Program, a nationwide program funded by the U.S. Office of Education to help selected low-income high school students prepare for and enter post secondary education. This paper describes the sample design and selection procedures used in selecting a sample of Upward Bound students and a sample of comparison students for the study. A multistage probability sample design was used. In addition, data collection procedures, which included questionnaire administration, mail queries and telephone followups, are described. (Author/BW)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED129835

TM005 233

RESEARCH TRIANGLE INSTITUTE

SAMPLE DESIGN AND DATA COLLECTION: NATIONAL STUDY OF THE UPWARD BOUND PROGRAM

by
Jane W. Bergsten
Senior Statistician
Sampling Research and Design Center
Research Triangle Institute

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Paper presented at the
1976 Annual Meeting of the
American Educational Research Association, Session No. 2902
San Francisco, California

23 April 1976

2

RESEARCH TRIANGLE PARK, NORTH CAROLINA 27709

ABSTRACT

The Sample Design and Data Collection Procedures:

National Study of the Upward Bound Program*

Dr. Jane W. Bergsten

Research Triangle Institute

The Research Triangle Institute conducted an evaluation study of the Upward Bound Program, a nationwide program funded by the U.S. Office of Education to help selected low-income high school students prepare for and enter post secondary education. This paper, the third in a series of four, describes the sample design and selection procedures used in selecting a sample of Upward Bound students and a sample of comparison students for the study. A multistage probability sample design was used. In addition, data collection procedures, which included questionnaire administration, mail queries and telephone followups, are described.

*Paper presented at the 1976 Annual Meeting of the American Educational Research Association, San Francisco, California, April 23, 1976.

SAMPLE DESIGN AND DATA COLLECTION: NATIONAL STUDY OF THE UPWARD BOUND PROGRAM

A. Sampling Overview

A major purpose of this study was to measure the effect of the Upward Bound (UB) program on the educational continuance rates of the high school students it serves.[1/] In order to measure this effect, it was necessary to have some standard against which to compare the rates of the Upward Bound participants. The study, therefore, also concerned itself with measuring educational continuance among students that were similar enough to the Upward Bound participants to provide meaningful benchmarks again which rates of Upward Bound participants could be compared.

Ideally, a researcher would set up an experimental design in order to measure the effectiveness of such a program as Upward Bound. Students would be identified as meeting the acceptance standards of the program, and those eligible would be randomly assigned to the actual program or to a control group. Both the program participants and the control group students would be observed for a number of years, and the continuance rates calculated. By comparing the continuance rates for the Upward Bound program participants to those of the control group, the effect of the program on educational continuance could be estimated. Such types of experiments are rarely feasible and in many cases are administratively and policically impossible. Often, as in this case, a program has to be evaluated after it has been in effect for a number of years. However the evaluation procedures available at that stage cannot provide the precise answers that would be forthcoming from an experimental design.

This study design, and the associated sample design, focused on the effect of the Upward Bound program on educational continuance rates, as

---

[1/] The three main objectives of the study were to measure the effect of the UB program on (a) the high school retention rates of its participants (b) the post secondary institution entry rates of its participants and (c) the generation of the skills and motovation necessary for success and education beyond high school.

measured during a short span in time, after the program had been in operation for a number of years. To obtain measures of program effect the study measured the continuance rates, over the course of a one year period, of (1) Upward Bound participants who were high school sophomores, juniors, or seniors at the beginning of the one year period and (2) comparison students in these same three grades.

Ideally, such rates as described above would be computed for the entire population of Upward Bound participants in the grades of interest and for the entire population of comparison students. However, collecting data on an entire population is often both administratively cumbersome and financially impossible. Since statistical procedures exist whereby unbiased estimates of population characteristics can be obtained from data based on only a sample from the population, rather than the entire population, a sample survey is often the best way to obtain the desired information.

A probability sample can be selected in such a way that, not only can unbiased estimates, such as number of tenth grade dropouts, be made from the sample data, but estimates of sampling errors can be calculated as well. Sampling error is error that must be tolerated because we choose to select and measure only a sample of elements rather than all elements in a population. The sampling error (or standard error) provides a measure of the range within which a sample estimate can be expected to fall a certain percentage of the time. The magnitude of the sampling error is related to two things over which the sampler can exert some control, namely, the size of the sample and the procedures used in selecting the sample.

In developing a sample design, the sampler is concerned with the selection of a sample that will yield estimates of sufficient precision,

i.e., estimates having a small enough sampling error so as to be useful, and with producing these estimates for the least cost. In most instances a sample that is widely spread over the entire population will produce the most precise estimates. However, the cost of conducting a study with such a widespread sample is usually much greater than the cost of a study using a cluster sample (i.e., a sample where several or many elements are selected from the same place). By balancing both expected sampling errors and expected costs, the most desirable sample design can be determined.

For the Upward Bound evaluation study, rather than selecting a sample of Upward Bound participants without regard to their location, the selection was made by first selecting a sample of Upward Bound projects. In determining the number of projects and the number of Upward Bound students to include in the sample, the goal was to produce the most precise estimates for the least cost.

Since the purpose of the study necessitated the estimation of continuance rates not only for Upward Bound students but also for the comparison group of similar students, a comparison population had to be defined and sampled. It was decided that a group that satisfied both the requirement of similarity and of comparability to the Upward Bound students was the population of students attending the same schools as the Upward Bound students. Choosing this definition of a comparison group had both analytical and administrative advantages. By using as comparison students those who attended the same schools as UB students, we were controlling the differential institutional effects. This can be expected to have the effect of increasing the precision of estimates of differences between UB students

and comparison students. In addition, using such schools provided for geographic clustering of the sample UB and the sample comparison students and an associated reduction in cost.

The proposed definition of comparison students has one obvious disadvantage. To the extent that UB activity in a school has had a beneficial effect on students who have in no way been connected with the UB program, the st results would be biased in the direction of underestimating the beneficial effect of the UB program on educational continuance rates. However, all things considered, the proposed definition was judged to be the most desirable and the comparison group of students was therefore defined as 10th, 11th, and 12th grade students who were not in a UB program but who were attending schools that were also attended by Upward Bound students. These schools will be referred to as "feeder" schools.[2/]

Considering the precision of the estimates to be made from the samples, the costs involved in conducting the study and the administrative feasibility of carrying out field procedures, it was decided that 54 of the 333 Upward Bound projects would be selected into the sample. All students who were Upward Bound participants in a sample project during September of October 1973 were included in the sample. In addition, for each sample project, two "feeder" schools were selected and from each "feeder" school six or seven comparison students were taken from each of grades 10, 11, and 12. To further control the amount of field work involved in the sampling procedures, it was decided to sample students within the selected schools in two stages, by first selecting a sample of classes and then selecting a sample of students within the selected classes.

[2/] For the schools actually selected into the sample, an average of only 1% of the 10th, 11th, and 12th graders were actually involved with a UB program. The spillover effect of the UB program on the non-UB students in the same school might actually have been minimal.

In addition to the student samples, two different sample of staff members were selected: (1) staff from Upward Bound projects, and (2) staff from "feeder" schools. In addition, a subsample from the sample of 54 projects was selected for site visitation. The procedures for the selection of the UB and comparison student samples will be described in some detail. The other sampling procedures will not be covered in this paper.

A graphical description showing the different stages of selection of the UB and comparison student samples is presented in Figure 1. The procedures used at each stage of selection are described in the following sections.

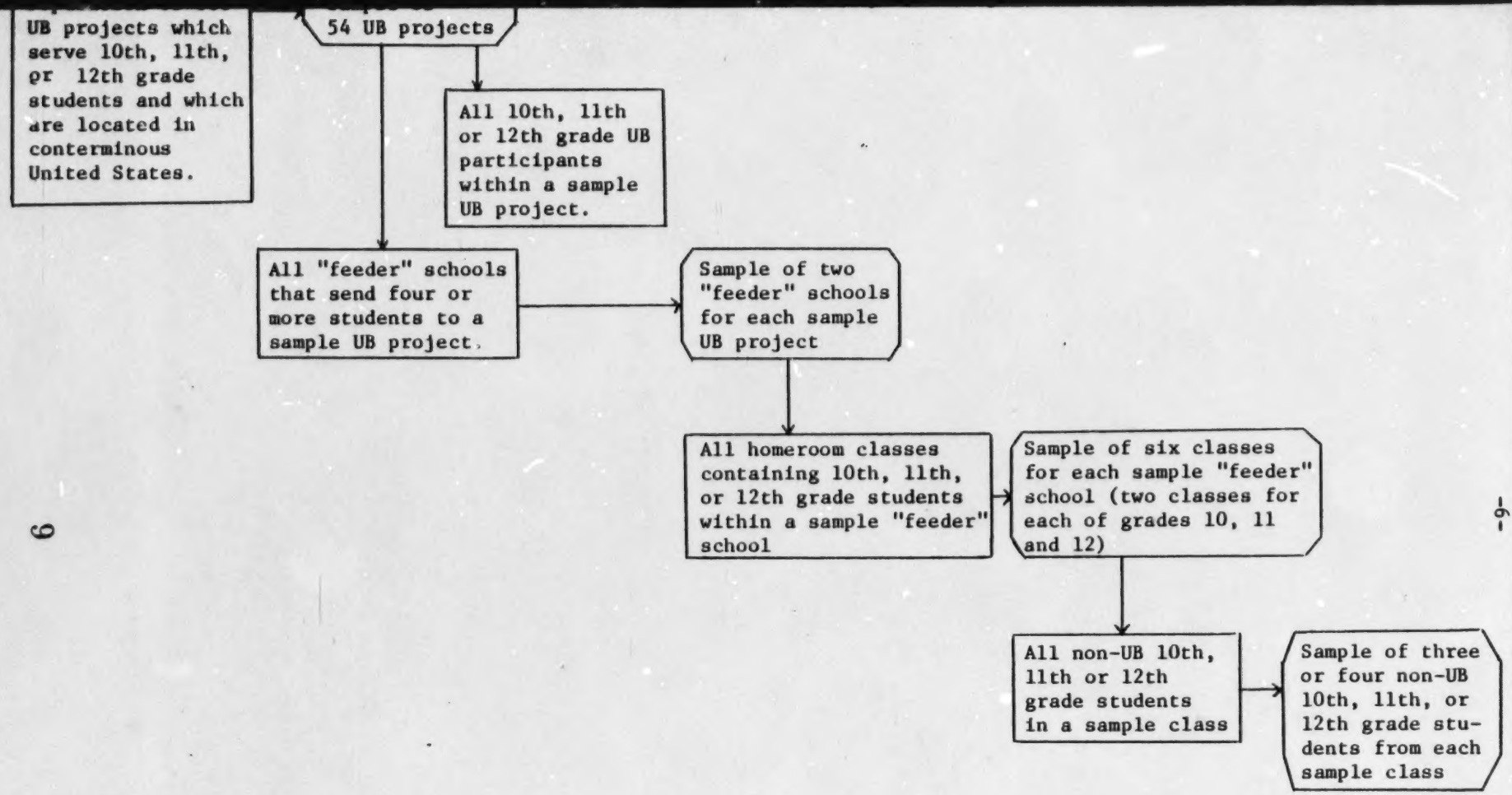
UB projects which serve 10th, 11th, or 12th grade students and which are located in conterminous United States.
54 UB projects
All 10th, 11th or 12th grade UB participants within a sample UB project.
All "feeder" schools that send four or more students to a sample UB project.
Sample of two "feeder" schools for each sample UB project
All homeroom classes containing 10th, 11th, or 12th grade students within a sample "feeder" school
Sample of six classes for each sample "feeder" school (two classes for each of grades 10, 11 and 12)
All non-UB 10th, 11th or 12th grade students in a sample class
Sample of three or four non-UB 10th, 11th, or 12th grade students from each sample class

Figure 1. The UB Student and Comparison Student Samples

B. Sampling Procedures[3/]

1. The UB Project Sample

Research Triangle Institute (RTI) obtained, from the National Office of Education (OE) and from the ten regional OE offices, copies of the project proposals for all UB projects that were funded for program year 1973-74. From these proposals characteristics of the funded projects were obtained, and a frame from which projects would be sampled was constructed. To have been included in the project sampling frame a project must have possessed all of the following characteristics:

a) The project was funded for fiscal year 1973.

b) The project proposed to serve students in grades 10, 11, and/or 12.

c) The project was located in coterminous United States.

The 333 projects possessing all of the above characteristics constituted the frame from which sample projects were selected. Using project characteristics obtained from the project proposals, the 333 projects were partitioned into 27 strata on the basis of six variables. These variables together with their code categories, are listed below.

a) Ethnic background of participants.

(1) Black.

(2) Black and white.

(3) 10%+ American Indian.

(4) 10%+ Mexican American.

(5) 10%+ Puerto Rican.

(6) All other.

[3/] For more detailed discussion of the sampling techniques described in this report, see Kish, L. (1965), Survey Sampling, New York: John Wiley & Sons.

b) Office of Education geographic region

(Region No. 1 through No. 10)

c) Program emphasis.

(1) Academic orientation.

(2) Vocational orientation.

(3) Combination of academic and vocational orientation.

(4) Other.

d) Project age.

(1) New this year.

(2) One year old.

(3) Two years old.

(4) Three or more years old.

e) Location of participants.

(1) Citywide

(2) Selected parts of a city.

(3) Rural.

(4) City and rural.

(5) Statewide.

(6) Reservation.

(7) Regional.

f) Type of sponsoring institution or agency.

(1) Secondary school.

(2) Two-year college.

(3) Four-year college.

(4) Vocational-technical school.

(5) Private, nonprofit agency.

(6) Consortium of education institutions.

(7) Agency for a consortium.

(8) Proprietary school.

(9) Public agency.

The strata were constructed so as to have approximately equal total number of participants in grades 10, 11, and 12. From each of the 27 approximately equal-size strata, two different projects were selected for the sample with probability proportional to a measure of size which reflected the total number of participants in grades 10, 11, and 12[4/]. This resulted in the selection of 54 different UB projects.

It may be helpful to consider the stratification and selection of the 54 sample projects in more detail. Stratification may be defined as the dividing of a population into sub-parts called strata, for the purpose of sampling separately from each stratum. Although there are a number of reasons for stratifying prior to the selection of the sample, the 27 strata were formed for the selection of 54 sample Upward Bound projects for two basic reasons: (1) to insure that certain minority groups such as American Indian, Mexican-American, and Puerto Rican would be properly represented so that there would be a sufficient number of such students in the sample to enable us to make relatively reliable estimates about these groups from the sample data; and (2) to attempt to reduce the size of sampling errors that otherwise would result if a sample of 54 projects were selected without stratification.

In forming the 27 strata, the aim was to make the projects within each stratum as similar as possible on the variables of central interest in the study, e.g., education continuance rate for tenth grade UB participants,

---

[4/] The procedure used is described in detail in M. R. Samford, "On Sampling Without Replacement with Unequal Probabilities of Selection," Biometrika, 54, 3 and 4 (1967), 499-512.

college entrance rate among twelfth grade UB participants, etc., and to make the 27 strata as different from one another as possible, on these variables. To the extent that this is done successfully, the sampling errors will be decreased, as compared to those based on an unstratified design.

Ideally one would form strata on the basis of the central variables themselves, but because such information is in general not available, the strata are usually formed on the basis of characteristics thought to be highly correlated with the central variables of the study. To the extent that stratifying on these related characteristics actually does form strata that are, in fact, homogenous with respect to the central variables under study, the sampling errors of the estimates of these variables will be reduced.

The stratification variables used in dividing the 333 projects into 27 strata were those thought to be related to school continuance rates: ethnic background, geographic region, Upward Bound project program-emphasis, project age, location of participants, and project sponsor. If all of these six variables were cross classified with one another, there would be some 108,000 cells. Since there are only 333 projects, most of the cells would be empty, and many would contain only one or two projects. It is obvious that considerable combining of cells was necessary. The combining process resulted in certain categorizations being used for some groups and not for others. For example, city versus non-city was a meaningful and useful way to divide projects having a sizeable number of Puerto Ricans, but was not useful in dividing projects classified as American Indian. For this latter group, a reservation versus non-reservation distinction was made.

In forming the homogeneous strata, an attempt was made to form the strata so that they were as nearly equal in size as possible (i.e., equal with respect to the number of Upward Bound participants anticipated for the school year 1973-74). This was done because, in general, using equal size strata can be expected to bring about greater gains in precision of estimates relating to student characteristics than using strata that very greatly in size.

Twenty-seven strata were used for the selection of the 54 sample projects, with two projects being selected from each stratum. This set of 27 strata, of course, form only one set from among the thousands of possible sets of 27 strata that could have been formed. It is entirely possible that another set would have been better, that is, would have yielded estimates having a somewhat smaller sampling error than estimates from a sample based on the 27 strata that were actually used. Undoubtedly many other sets would have been worse, that is, would have provided estimates with somewhat larger sampling errors. Different ways of forming 27 strata might affect the size of the sampling errors of the estimates, but in no way would the unbiasedness of the estimates that can be made from the study data be affected. Unbiased estimates of population characteristics are possible regardless of the effectiveness of the stratification.

Once the 27 strata were formed, 54 projects were randomly selected, two from each stratum using probabilities proportional to a measure of size and without replacement. By selecting two projects from each stratum it is possible to make unbiased estimates of error variances.

2. The Student Samples

a. The UB Student Sample

Within each of the 54 sample projects, all students who were in

UB participants in September or October of 1973 were selected into the sample. This resulted in the selection of the 3710 eligible UB students associated with the 54 sample UB projects.

b. The Feeder School Sample

A sample of control or comparison students against which to compare the UB students was selected. The comparison students as well as certain staff members were selected from "feeder" schools associated with the 54 sample projects. A feeder school was defined as a school with at least four 10th, 11th, and/or 12th grade students listed on the most recent CUB[5/] roster as being clients of a given UB project, or as a school from which a newly funded UB project planned to serve at least four 10th, 11th, or 12th grade students during the 1973-74 school year.

For sampling purposes each "feeder" school was associated with one and only one UB project. If a "feeder" school "fed" more than one project, it was associated only with that project to which it provided the greatest number of students. In order to accomplish this procedure of uniquely associating a "feeder" school with one and only one UB project, the CUB roster which listed the schools attended by UB students was checked for all 333 projects in the defined population.

Once the "feeder" schools for a sample project had been identified, they were listed in descending order according to the number of UB students who attended the school and who were associated with the sample Upward Bound project. The estimated total number of students in grades 10 through 12 was then recorded for each "feeder" school, and these numbers were then accumulated. A systematic random selection of four "feeder" schools was

---

5/ The Current Upward Bound Student Roster that was correct as of August 1973.

then made with probability proportional to size. An equal probability sub-selection of two of the four schools was then made and designated as "sample" schools. The remaining two schools were designated as "backup" schools to be used only if the "sample" schools refused to participate in the study. For some projects additional "feeder" schools were selected. This was done to keep the comparison student weights from being excessively large.

A total of 113 school selections were made, bringing into the sample 108 different schools. The selection procedures permitted a school to be selected more than once, and in five instances a school was selected twice. In such cases, the number of sample students to be selected from the school was doubled.

Of the 108 sample schools selected, five refused to participate. Four "backup" schools were selected as replacements, which brought the total to 107 schools that were expected to participate. The late refusal of one school resulted in a sample of 106 schools providing the sample comparison students that were used in the study. (See Table 1)

c. The Comparison Student Sample

Within each sample "feeder" school, six homeroom classes of students in grades 10 through 12 were selected with equal probability from class lists supplied by the sample schools. For each selected class a list of students who were in the class during September 1973 was obtained. For each student listed, the homeroom teacher was asked to indicate the student's grade level and ethnic background, and to make an "educated guess" as to whether or not the student came from a low income family and as to whether or not the student should be considered as an "academic risk" for a two- or four-year college education. This information was used to

Table 1

Distribution of sample feeder schools by study participation

	Number of schools in sample	Number of school selections in sample*
Number selected into sample	108	113*
Number refusing to participate in study	5	5
Number of back up schools selected as replacements for refusals.	4	5**
Total in adjusted sample	107	113
Number refusing to participate in study (during data collection stage)	1	1
Total number participating in study	106	112

*The sampling procedures permitted a school to be selected twice, in which case the number of students to be selected from the school was doubled. The 113 school selections resulted in the selection of 108 different schools, 5 of which had been selected twice.

**To adjust for the 5 schools that had refused, 4 "back up" schools were selected to replace 4 of the 5 refusals. In the case of the fifth refusal, the sample size in the cooperating sample school associated with the sample project was doubled.

stratify students prior to selection. (Rather than sending class lists, several sample schools sent complete rosters of all of their students who had been enrolled in September 1973. In such cases a sample of 20 students from each grade level was selected, and the lists of selected students were then sent back to the school with the request that the school provide information on the background characteristics of each student listed. Such completed lists were then treated as though they were class lists, for the purposes of selecting the sample comparison students for the study.)

The student lists for the sample classes were carefully checked against the September 1973 list of UB students, and the students who were listed as being associated with an Upward Bound project were removed from the student class lists. From these corrected lists, a stratified random sample of comparison students was then selected, with an expected 21 students from each sample control school, yielding a total of 2340 eligible comparison students.

The selection of the approximately 21 students per school involved a predetermined set of procedures. First, using the information provided by the homeroom teacher, the students were stratified according to grade-in-school and according to whether or not they appeared to meet the eligibility requirements of the UB program. Then, a random sample of students was selected from each stratum, with the objective of selecting approximately equal numbers of students from each of the three grades, and of over-sampling students who <u>appeared</u> to be like UB students. While we thus selected into the sample students tentatively classified as "like UB" as well as those tentatively classified as "unlike UB" students, we included in our sample a disproportionately high number of "like-UB" students.

This was done in order to attempt to insure the inclusion of a substantial number of students who were, in fact, eligible for an UB program. (In the analysis of the study data, the characteristics of comparison students were, of course, determined from different factual information that the teachers "educated guesses" that were used in the sample selection procedures.)

The approximate relative sampling rates for students in each of four income academic risk groups is presented in Table 2.

Table 7
Approximate Rates of Selecting Comparison Students

	Family Income Not Classified as "Low"	Low Family Income
Academic risk	1 x rate	4 x rate
Not Academic risk	1 x rate	2 x rate

Note that "low income-high academic risk" students were selected at approximately four times the base rate and other low income students were selected at approximately twice the base rate.

At each stage of the sample selection, each unit in the defined population was given a positive chance of being selected. These probabilities were recorded and combined into an overall probability of selection for each UB and each comparison student, thus providing for the computation of unbiased estimates of population characteristics for UB students as a whole as well as for comparison students as a whole. In order to produce unbiased estimates, each student in each of the two samples was ultimately assigned a weight equal to the inverse of the probability of his being selected. Thus, students selected with smaller probabilities received larger weights, and vice versa. The weighting procedures used at the estimation stage

compensated for unequal probabilities having been used at the selection stage, and permited the unbiased estimation of characteristics of the population of UB students and of the population of comparison students. In addition to making unbiased population estimates the study plan called for making comparisons of the two groups of students, the UB and comparison students. This latter population, however, was not necessarily expected to be similar to UB students on all revelant background characteristics. In order to reduce or eliminate the effects of these dissimilarities, a balancing or standardizing statistical technique was planned and ultimately employed. This technique, which will be described in some detail, statistically adjusted the comparison student population to the UB population using techniques similar to those employed by demographers when they construct "adjusted", "standardized", or "corrected" birth rates or death rates. The aspect of the sample design related to the oversampling of comparison students thought to be "like" UB students aimed at increasing the precision of the comparisons between UB students and the "balanced" or "standardized" comparison student population. The balancing procedures and estimation techniques will be described in more detail in a later section.

C. Data Collection Procedures

Although the populations to be sampled were defined according to their UB participation and school attendance status as of September or October, 1973, the actual data collection did not begin until the spring of 1974, at which time the sample UB and sample comparison students were queried.

Prior to the spring data collection initial contacts had been made with the UB project directors and with state and local school officials in areas where sample schools were located in order to secure their endorsement. Sample school personnel were then contacted to obtain information for sampling purposes and to pave the way for the data collection procedures. During the initial contacts with a sample school a "school contact" individual was appointed to serve as a laison between RTI and the school. The school class lists, the student lists for the sample classes, etc. were obtained through the "school contact".

Sixty-two study administrators were hired to conduct questionnaire administrations and to collect school transcript data at the 54 UB projects and 107 sample feeder schools. Six regional training sessions were held to provide the study administrators with the study objectives, and to familiarize them with the forms and procedures to be used in carrying out the questionnaire administrations, obtaining transcript information, etc.

The Basic Student Questionnaires (BSQ) were administered to UB students (1) in group sessions during regularly scheduled meetings of full UB membership at a sample project and (2) in remote site administrations where projects held only infrequent meetings or where the meetings were poorly attended. In addition, the study administrators collected basic transcript information

from the project files. The BSQ were administered to comparison students in group sessions held at the sample feeder schools, and transcript information was obtained for sample comparison students from school files by the study administrator. For both UB and comparison students, make-up sessions were held for those students who did not show up for the initial questionnaire administration sessions.

Drop-out-Transfer Questionnaires were mailed to "absentee" sample UB and comparison students, and an additional mail follow-up was sent to non-respondents.

There was great concern on the part of RTI staff members to insure a high response rate, and steps such as scheduling extra make up sessions during UB summer programs and mailing BSQ questionnaires were taken.

The difficulty encountered in obtaining a high response rate during the spring data collection caused some concern about the prospects for the fall status questionnaire that would be mailed out during the early fall of 1974. A small pilot study was conducted during early September 1974 to determine the feasibility of using a telephone interview to obtain the desired survey information. The results of pilot study convinced us that the FSQ response rates would undoubtedly be low, and that they would be especially low among previous non-respondents.

The mail returns of the fall status questionnaire did indeed fall far below what had originally been expected. The low response rate on the FSQ caused more than the usual amount of concern because it was felt that whether or not the sample UB or comparison student responded might well be highly related to whether or not the person was still enrolled in school. Such a relationship could bring about spurious results if estimates were based on approximately 50 percent responding.

It was reasoned that a UB or comparison student who had dropped out of school would be more likely to have left home and less likely to have received and responded to the mail FSQ. If such were the case, the school continuance rate among non-respondents could be far less than the continuance rate among respondents to the FSQ.

Special steps were taken to obtain fall status information from those UB and comparison students not responding to the FSQ. A telephone tracing procedure was employed wherein the sample UB or comparison student was telephoned and the desired information was obtained for the sample individual by means of a telephone interview. In cases where the individual himself could not be contacted, the desired information was sought from other knowledgable individuals such as parents, project directors, school personnel, etc.

Although it would have been preferable to contact all non-respondents to the FSQ by phone, cost considerations made this prohibitive. Instead, a sampling plan was instituted. The most worrysome groups were those UB and comparison students from whom no response to any questionnaire had yet been obtained. All such students, UB and comparison (exclusive of refusals) were subselected with certainty into the telephone tracing subsample. For those students, not only was the FSQ information collected by phone, but also certain pertinant information that should have been collected during the previous spring. Of the remaining non-respondents to the FSQ, all of whom had responded to a previous query, approximately 40 percent were randomly selected for inclusion in the telephone tracing subsample. Compensation for the subselection procedure would be made prior to estimation by a special weighting procedure.

The sample sizes and responses received are detailed in table 3.

Table 3

Instrument Response Rates

Group	Instrument	Number Eligible	Percent return for all eligibles
	SPRING 1974 DATA COLLECTION		
Upward Bound	Basic student questionnaire	3,337	82.8%
	Dropout transfer questionnaire	373	37.8%
		3,710	
	Student transcript form	3,710	100.0%
Comparison	Basic student questionnaire	2,082	85.1%
	Dropout transfer questionnaire	258	25.6%
		2,340	
	Student transcript form	2,340	99.1%
	FALL 1974 DATA COLLECTION		
Upward Bound	Fall status questionnaire		
	Mail: previous respondents	3,179	62.8%
	previous nonrespondents	531	22.0%
		3,710	
	Telephone: previous respondents	1,183	38.6%*
	previous nonrespondents	414	97.8%
Comparison	Fall status questionnaire		
	Mail: previous respondents	1,838	58.3%
	previous nonrespondents	502	15.5%
		2,340	
	Telephone: previous respondents	767	38.3%*
	previous nonrespondents	424	90.6%

*Reflects primarily subsampling. Response rate among those sampled was approximately 95%.

D. Estimation

Since all of the samples selected for this study are probability samples, it is possible to make unbiased estimates of population characteristics from properly weighted sample data. The weights used in making these estimates are determined from the probabilities of selection. Each element in the sample was assigned a weight equal to the inverse of the probability of its being selected in the sample.

For example, to obtain weights for comparison students, we first determined the overall probability of selection for each comparison student. These were determined as follows:

[erall probability of lection of mparison udent] = [Probability of selecting the UB project associated with his school] · [Probability of selecting his school, given that the associated project had been selected] · [Probability of selecting his class, given that his school had been selected] · [Probability of selecting the student, given that his class had been selected]

The weight assigned to a student was then determined by taking the inverse of his overall probability of being selected.

$$\text{(Student weight)} = \frac{1}{\text{(overall probability of selecting the student)}}$$

Each UB participant, comparison student, (and staff member) selected into the sample was assigned a weight which was computed from his overall probability of selection. (In addition, each of the 54 projects was assigned a weight, and each of the 15 projects subselected for site visitation was assigned a second weight.) All population estimates made from the sample

data used these sample weights.[6/]

If responses were available for every sample individual in a specified subgroup, for example UB students, an estimated total could be obtained by merely summing the weights of the sample individuals in that subgroup.

Because there was some non-response, a special weight adjustment was applied to responding sample students, to compensate the nonresponses. Non-responses adjustments were made within homogeneous categories, such as grade-project categories for UB students, and project-grade-balancing group[7/] categories for comparison students.

[6/] Detailed formulas for estimating population totals, means, and proportions and their respective sampling errors are presented in: Graham J. Burkheimer, et.al. Evaluation Study of the Upward Bound Program (Volume IV of A Study of the National Upward Bound and Talent Search Programs. Four Volumes) Research Triangle Park, North Carolina: Center for Educational Research and Evaluation, Research Triangle Institute, March 1976.

[7/] The balancing groups, which were formed on the basis of ethnicity, sex, poverty status, and academic risk status, are defined more completely in a later section.

E. Standardization or Balancing

The population of comparison students, i.e., the group against which the UB students were to be compared, was defined as all non-UB students enrolled in the 10th, 11th or 12th grade of an UB project feeder school[8/] in September 1973. There was no reason to believe that the distribution of the population of comparison students on relevant background characteristics would be exactly the same as those of the UB population of students. Therefore, any direct comparisons between UB students and comparison students with respect to such variables as school continuance rates would not take account of the fact that the distributions of the groups being compared might be different on characteristics such as family income, ethnic background, etc. An adjustment for such differences in these background characteristics was made by using procedures similar to those used by demographers in constructing what are referred to as "adjusted," "standardized," or "corrected" birth rates or death rates.[9/] The demographer, for example, computes age specific birth rates for each relevant age group, and applies them to a population of interest, age group by age group, in order to obtain an "age adjusted" birth rate. In a similar fashion we could, for example, compute school continuance rates for each of several family income categories, for 10th grade comparison students. By applying these rates to a population of 10th grade comparison students having the same income distribution as the UB student population, an "adjusted continuance" rate could then be computed for comparison 10th grade students. The comparison of the continuance rate for UB students with the "adjusted" continuance rate for comparison students would be free of the effects of differential family income distributions among UB 10th graders as compared to comparison 10th graders. By computing such "adjusted" rates the effects of differences in background characteristics of the UB students and the comparison students could be reduced or eliminated.

[8/] For definition of feeder school, see Section A.

[9/] See Spiegelman, M. Introduction to Demography. Harvard University Press, 1968, Cambridge, Massachusetts. Sections 4.3.0-4.3.1.

A hypothetical example is presented in Table 4 to show how different background characteristics could affect the results of comparing two groups such as UB and CS, and how a balancing or standarization procedure could be employed. In Table 4 the proportion of students in each of the three racial groups is presented for UB and CS students. In addition, a fictitious "average-scale-score" is given for each of the six race by type-of-student subgroups. Note that when the average-scale-scores for UB and CS students are compared within each of the three racial groups, it is clear that the UB students have lower average scores. Consistently, in each of the three racial groups, the average scale score for UB students is 10 points lower than for CS students. For Blacks the UB to CS comparison is 100 to 110, for whites it is 70 to 80 and for "other" it is 110 to 120. Note, however, that the average-scale-scores for all races combined show a reverse relationship, 97 for UB as compared to 95 for CS. The reversal in the relationship is brought about by

Table 4

HYPOTHETICAL EXAMPLE OF BALANCING TECHNIQUE

Race	Upward Bound Proportion as students	Upward Bound Average scale score	Comparison Proportion of students	Comparison Average scale score	Balanced Comparison Proportion of students	Balanced Comparison Average scale score
Blacks	.61	100	.28	110	.61	110
White	.18	70	.55	80	.18	80
Other	.21	110	.17	120	.21	120
All races	1.00	97*	1.00	95**	1.00	107***

*The average-scale-score is obtained as follows:
(.61)(100)+(.18)(70)+(.21)(110)=97

**The average-scale-score is obtained as follows:
(.28)(110)+(.55)(80)+(.17)(120)=95

***The average-scale-score is obtained as follows:
(.61)(110)+(.18)(80)+(.21)(120)=107

the differential racial distributions of the UB and CS groups. The differences in racial makeup of the two groups have obscured the true situation. A balancing technique which applies the average-scale-scores of the CS group to a CS population with the same racial distribution as the UB groups yields an average scale score of 107 for CS. This permits a comparison of UB and CS average-scale-scores that is free of the effects of differential racial distributions. Note that the UB-CS comparison on average-scale-score, after balancing, is 97 vs. 107. This reflects the 10 point difference in scores that appeared in each of the three racial groups.

While the actual balancing technique that was used in the analysis involved more than three balancing categories, the example shows the general way in which the balancing technique was applied.

The procedures actually used involved standardizing or balancing the CS group of students to the UB population characteristics within each of the three grades 10-12. For each of grades 10, 11 or 12 separately, 16 balancing groups were formed for UB students, and again for CS students. The 16 balancing groups were based on two variables associated with UB participation qualifications, namely poverty status and academic risk, and on two additional variables, ethnicity and sex. The 16 balancing groups were formed as follows:

Table 5
Definition of Balancing Groups

	Poverty Status and Academic Risk			
Ethnicity and Sex	Poverty Risk	Poverty Not "Risk"	Not "Poverty" Risk	Not "Poverty" Not "Risk"
	Balancing Group Number			
Black, male	1	2	3	4
Black, not "male"	5	6	7	8
White, all sexes	9	10	11	12
Other, all sexes	13	14	15	16

Within each of the three grades separately, for each of the 16 balancing groups, the total estimated population size was obtained for UB students and for comparison students. Using these numbers a balancing weight component was computed for each comparison student. Each comparison student's final sample weight[10/] was then multiplied by his balancing weight component to obtain his final balanced sample weight.

The computations were thus carried out as follows:

For a given grade, g=10, 11 or 12 let

$Y_{ug}(k)$ = the estimated number of UB students in grade g in balancing group k.

$Y_{cg}(k)$ = the estimated number of CS students in grade g in balancing group k.

$$\omega_{ug}(k) = \frac{Y_{ug}(k)}{\sum_{k=1}^{16} Y_{ug}(k)}$$ = proportion of grade g UB population in balancing group k

$$\omega_{cg}(k) = \frac{Y_{cg}(k)}{\sum_{k=1}^{16} Y_{cg}(k)}$$ = proportion of grade g CS population in balancing group k.

The balancing weight component for each comparison student in grade g in balancing group k is then

$$W_{cg}^{\beta}(k) = \frac{\omega_{ug}(k)}{\omega_{cg}(k)}$$

By multiplying his final sample weight by $W_{cg}^{\beta}(k)$, the comparisons students final balanced weight was obtained.

Balanced or standardized rates, proportions, etc. could then be estimated for the comparison students in a given grade merely by using the final balanced weights rather than the final sample weights when applying the estimation formulas. The comparing of UB rates, proportions, etc. with balanced CS rates, proportions, etc. would

---

10/ The student's final sample weight consisted of the product of two weight components. The first weight component was the inverse of the student's probability of selection into the sample and the second was a component used to compensate for nonresponse.

then be free of the effects of differential distributions among the balancing groups.[11/]

[11/] Detailed information for estimating balanced means and proportions, differences between UB means and proportions and comparison student balanced means and proportions, and their respective sampling errors are contained in the previously cited detailed report.